Optimal Pricing of a New Utility Service : The Case of Piped Water in Vietnam*

Quy-Toan Do[†] Hanan G. Jacoby[‡]

Abstract

As utility services expand throughout the developing world, providers must grapple with how to set prices to recover average costs. Data from a multi-year randomized pricing experiment among nearly 1500 recently-connected piped water customers in Vietnam reveal month-to-month demand persistence. Based on structural demand estimation, we document how endogenous preferences, if unaccounted for, can lead to low take-up and thereby threaten the financial viability of the new water utility. We also show that such demand persistence calls for pricing schemes that defer lump-sum payment, effectively allowing future consumers to subsidize their present selves.

JEL codes: L95, D91, O18

Keywords: Endogenous preferences, two-part tariffs, willingness-to-pay.

*This study is registered in the AEA RCT Registry and the digital object identifier (DOI) is: "10.1257/rct.6589-1.0". We are very grateful to the Editor Francesca Molinari and to three anonymous referees for comments and suggestions that greatly improved the paper. We also benefited from discussions with Jishnu Das, Alejandro Molnar, Claudia Ruiz, and Gil Shapira and from comments by seminar participants at the World Bank applied-micro seminar. We are especially thankful to Bruce Hansen for econometric advice. Financial support from the World Bank's Strategic Impact Evaluation Fund (SIEF), the Knowledge for Change Program (KCP), the Impact Evaluation to Development Impact (i2i) Fund, and the Research Support Budget is gratefully acknowledged. Tung Duc Phung and the Mekong Development Research Institute (MDRI) managed the price experiment and collected the baseline data. Trung Dang Le and RTAnalytics (RTA) collected quarterly monitoring data. Irène Hu provided excellent research assistance. The findings, interpretations, and conclusions expressed in this work do not necessarily reflect the views of the World Bank, its Board of Executive Directors, or the governments they represent. The World Bank does not guarantee the accuracy of the data included in this work.

[†]World Bank: qdo@worldbank.org

[‡]World Bank: hjacoby@worldbank.org

1 Introduction

The provision and pricing of utilities remain major policy issues in developing countries, where the expansion of piped water, electricity, and other large service networks is ongoing. Economists have long been concerned with how such natural monopolies should be priced, subsidized and/or regulated to ensure access to socially valuable services while covering their large capital costs (see, e.g., Laffont and Tirole 1993; Joskow 2007). A common premise in this literature is that consumer preferences are exogenous and unchanging. Once connected, however, it may take time for consumers to habituate to the new service or to learn about their preferences. As a result, current consumption may shift out future demand. If such a dynamic is sufficiently strong, and consumers are unaware of it prior to connecting, then they may substantially undervalue a connection ex-ante, leading to sub-optimal take-up and even non-viability of an ex-post socially desirable utility service. In this setting, an informed policy-maker could improve welfare by pricing a new utility so as to account for endogenous preferences and consumers' lack of awareness thereof.

To quantify the endogenous formation of preferences for a new utility service, we ran a three-year price experiment in a rural commune of Vietnam's Red River Delta Region served by a single piped water provider. Our setting is fitting because households were in the process of transitioning from traditional rainwater harvesting to modern piped water, both of which systems were installed in their homes concurrently. Rainwater needs to be stored in and pumped from dedicated tanks, as well as filtered and boiled for cooking purposes, whereas piped water provides greater convenience on these dimensions as well as higher pressure for showers, house cleaning, and the like. Reflecting the transition process, in the three years from the start of the water utility's operations to the start of our experiment, piped water use per customer grew at an average rate of 0.9 percent per month by volume, pointing toward a behavioral model that delivers at least an initial phase of non-stationary consumption. While our focus on piped water in Vietnam is, of course, particular, we believe that learning through experience and/or habit formation are generic features of newly introduced utility services in a wide range of settings, both contemporary and historical.[1]

[1]For instance, Taylor and Trentmann (2011) vividly recount how bathing habits and routines in

For the experiment, about 1,500 piped water customers were randomly divided into three equal groups. We then rolled out a subsidy scheme sequentially, with each group receiving a 50 percent discount on the price of piped water for six months, followed by a 75 percent discount for another six months after that. At any point in time, therefore, at least one group did not receive a subsidy and could thus serve as a control. To make the changes in price throughout the experiment immediately salient, we undertook a personalized communication campaign, both at home and at the time of monthly bill payment, which was always done in person at a central location. Meanwhile, we collected data from the utility on monthly piped water usage and, at a quarterly frequency, from households on, among other things, the number and duration of stay of residents (including visitors) and on water-sharing practices.

In addition to a modest response to the current price discount, the experimental data reveal persistence: households that faced a discounted water price during months in the recent past and thus had higher usage in those months, also consumed (modestly) more in the current month *conditional* on the current month's discount. This evidence, which exploits random assignment of discounts throughout the course of the experiment,[2] suggests that households were building consumption capital, either in the form of habits or learning through experience.[3] Alternative explanations can be plausibly ruled out: lags in adjustment to new prices were obviated by our proactive information campaign; loss aversion or reference dependence, which could lead to apparent demand persistence, is inconsistent with the symmetric demand responses to upward and downward price changes observed in the data and cannot account for demand growth prior to the experiment; finally, household investments in piped water infrastructure (such as plumbing and washing machines) cannot be the source of the experimental findings given the size and duration of the price subsidies.[4]

Victorian England evolved with both changes in the availability and in the pricing of piped water. More recently, Fobi et al. (2018) used longitudinal data from Kenya to document patterns of consumption growth since connection similar to ours in the context of residential electricity.

[2]As we will describe, while our experimental design also afforded a 'nonparametric' test for persistence by comparing one month of piped water consumption data across groups of households that received different subsidies in the prior 12 months, this test proved inconclusive.

[3]Dupas (2014) interprets experimental evidence on demand persistence for bednets in Kenya through the lens of a structural model to demonstrate a large experience good effect. Habit formation is a less plausible source of persistence in her setting.

[4]If household investments were indeed a key source of observed demand persistence and if households were unaware that they would be making such investments in the future, then they would also

To quantify the welfare gains from utility pricing strategies that leverage such persistence, we next use our experimental data to structurally estimate the dynamics of piped water demand. Under our maintained assumption that, in deciding on their current piped water usage, consumers do not internalize its effect on future demand, consumption depends only on the current price and on past usage.[5] Our experiment, by generating random variation in price, not only allows us to identify the (short-run) price elasticity of piped water demand but also delivers valid instruments, namely lagged prices, for causally estimating the effect of past consumption on current consumption. These dynamics imply a long-run elasticity of 0.126, nearly three times greater than the short-run elasticity of 0.046. Before turning to pricing simulations based on these estimates, we show that model predictions match the dynamic pattern of piped water consumption out-of-sample. Specifically, we start each household off with zero past consumption at the time they connect and use our estimated model to forecast their consumption trajectories from the month of the connection up until the start of the price experiment in January 2016. Our forecasts fit the actual out-of-sample consumption trajectories reasonably well.

We next extend the paradigmatic optimal utility pricing framework to the case of endogenous preferences. In the seminal treatment of two-part tariffs (Auerbach and Pellechio, 1978), a utility, required to cover its average cost, chooses a lump-sum connection fee and a marginal price to maximize a static welfare function defined over heterogeneous consumers who are free not to connect.[6] The optimal price for the utility service in this scenario may entail a markup over marginal cost, a *tax* disproportionately falling on high-demand consumers to subsidize connection fees and thereby draw in low-demand consumers. When the demand for the utility service grows endogenously and consumers are unaware of this ex-ante, however, there is also an incentive

underestimate their willingness-to-pay for piped water ex-ante. While the broad messages of this paper would thus still go through, the particulars of optimal pricing would differ, as would the welfare implications. Specifically, a model of the endogenous choice of durable goods complementary to piped water would have to be estimated.

[5] We justify this assumption of consumer lack of awareness in detail below as, among other things, consistent with projection bias (Loewenstein et al., 2003).

[6] Increasing block tariffs, whereby the marginal price increases with usage, may be more efficient if the price elasticity of demand is decreasing in usage (see the discussion of "Ramsey pricing" in Wilson, 1993).

to *subsidize* the price through a negative markup, which increases long-run welfare.[7] The net effect of these conflicting tax and subsidy motives on markup is, in general, uncertain. Further, because consumers' ex-ante valuations of the service, which govern their connection decisions, are lower than their valuations ex-post, i.e., after consumption reaches its long-run steady state, the utility has an incentive to *defer* lump-sum charges; this might take the form of replacing the upfront connection fee with a recurring subscription fee (as exist in most cell-phone plans, for example). Back-loaded lump-sum service fees are thus a novel consequence of endogenous preference formation. Of course, deferment of the connection fee could also be a way of providing credit to customers limited in their ability to pay upfront. But, even with credit, consumers may still be unwilling to take up the service, whereas their future selves would. Deferred payment thus effectively taxes the high-willingness-to-pay future-self to subsidize the low-willingness-to-pay present-self and, in doing so, reduces the need for a positive markup to redistribute from high to low-demand consumers.

The final step is to numerically solve for optimal water pricing schemes given estimated (endogenous) preferences. We first consider a scenario in which the utility incorrectly assumes static or exogenous preferences, as might be obtained from an ex-ante willingness-to-pay elicitation (e.g., Lee et al. 2020). In this case, the optimal connection fee and marginal price combination from the utility's point of view is just the Auerbach and Pellechio (1978) two-part tariff. Next, we recalculate the two-part tariff when the utility correctly accounts for the endogenous formation of preferences as well as their customers' lack of awareness thereof. We find that substantial welfare gains can be achieved by the "sophisticated" utility, depending on its costs. Given the average cost in our setting, pricing that ignores demand persistence would fail to generate enough revenue to finance the utility in the first place, whereas a sophisticated two-part tariff *would* cover the cost of this ex-post socially desirable project thereby leading to an average welfare gain of US$28 per customer per month. Lastly, we compute the optimal (deferred) subscription fee, which, by allowing for a lower marginal price, leads to an additional welfare gain of US$5.6 per customer per month over the

[7] A related literature on the monopoly pricing of experience goods shows that low initial prices followed by higher prices later may be optimal in some circumstances (Shapiro, 1983; Bergemann and Välimäki, 2006).

connection fee contract.

This study contributes to a growing body of evidence, from both developed and developing countries, of persistence and adjustment lags in the demand for utility services, whether in response to price changes (Jessoe and Rapson, 2014; Ito et al., 2018; Ito and Zhang, 2020; Deryugina et al., 2020) or to conservation programs (Allcott and Rogers, 2014; Costa and Gerard, 2018). Two features of our work set it apart, however. First, we focus on the case of a new utility service, in which the policy-maker faces a tradeoff between expanding the customer base and extracting more from established customers. Second, while these existing studies also rationalize their empirical findings by appealing to some form of habit formation and/or learning, they are not designed to structurally estimate endogenous preferences. In doing precisely this, we are uniquely able to make quantitative welfare statements about counterfactual pricing structures that incorporate such preferences.[8]

More broadly, in randomizing a relevant policy instrument, here the marginal price of the service, across a real population of utility customers, this research advances the "lab-in-the-field" approach to questions of utility provision and optimal pricing. In pioneering work, Lee et al. (2020) randomize electricity service connection fees across rural households in Kenya to trace out the extensive margin demand curve and compare it to the average cost curve for connections to decide the social desirability of the service. Our paper cautions that such an ex-ante willingness-to-pay elicitation alone is insufficient for making both optimal provision and pricing decisions when preferences for the utility service form endogenously. In particular, an ex-ante elicitation measures consumer surplus, not as a *function* of the marginal price, but only at an implicitly assumed (and unknown) value thereof. We show, however, that the determination of optimal provision and pricing requires both the short-run and long-run price elasticity of demand as well, both of which we estimate using our experimental data. Our approach also highlights the wedge between ex-ante and ex-post valuations of the utility service, a source of low take-up that is complementary to but distinct from credit constraints (Devoto et al., 2012b; Berkouwer and Dean, 2021).

[8] In this respect, our paper extends the static demand literature focusing on how utility price policy affects consumer welfare; see, e.g., Reiss and White (2005), Ito (2014), McRae (2015), and Szabó (2015).

The rest of the paper is organized as follows: We describe the experiment, data and demand persistence results in section 2. In section 3, we set up a generic model of consumer demand with intertemporally dependent preferences and, in section 4, present the structural estimation of this model. With these results in hand, in Section 5, we assess optimal pricing schemes for new utility services, and specifically for piped water in our setting. We conclude the paper in Section 6.

2 Experiment and data

Working with the Government of Vietnam, the overall objective of the experiment was to assess willingness-to-pay for piped water in a rural area where the service had recently been introduced and had begun supplanting the traditional mode of water supply, rainwater collection.

2.1 Experimental design

The price experiment took place between January 2016-May 2018 in three villages of My Huong commune in the Red River Delta region of northern Vietnam. My Huong is relatively prosperous, with average per capita expenditures at about the 80th percentile for rural Vietnam. Since the arrival of piped water in 2012, My Huong has been served by a single provider, An Thinh utility, with whom we partnered over the course of the 3-year project. Among other things, An Thinh shared their customer billing records, which include monthly water usage (in m^3) and payment, from the month of connection onward, including throughout the price experiment.

Survey instruments and sample

A door-to-door census of the three villages was completed in mid-2015, yielding a total of 1660 households, including multi-family living arrangements with shared kitchen and electrical meter. After matching with An Thinh's customer records, we verified that 267 households or 16 percent of the population were unconnected to piped water as of July 2015, at which time we offered them free connections (i.e., waived the connection fee). While almost all of these 267 households eventually did connect, only about half did so in time for the baseline survey.

We undertook a baseline survey in October 2015. Out of the universe of 1,526 connected households in that month, including some of those that we had offered free connections to in July, a total of 1,488 participated in the survey. The baseline questionnaire asked about household demographics, assets, income, consumption, and domestic water use. We also began a series of quarterly "monitoring" surveys in November 2015, implemented at the time of bill payment, which collected information on household demographics (whether households had family or friends returning/staying with them and for how many days) as well as water sharing practices (whether households were giving/receiving piped water to/from other households).

Our final experimental sample consists of 1,462 water customers.[9] Of these, 114 or 8 percent had been unconnected as of July 2015, but subsequently received a (free) piped water connection through the project in time to be included in the baseline survey. As noted above, prior to our interventions, a larger fraction of commune households (16 percent) did not have connections, presumably because they were not willing to pay the fee. Later, to compute social welfare in our pricing simulations, we augment the estimation sample by randomly drawing from the 114 households so that previously unconnected households (inclusive of the original 114) make up precisely 16 percent of the sample.

Water price interventions

The official pre-intervention tariff schedule for piped water had three increasing blocks: US$0.236 per m^3 for consumption below 10 m^3, US$0.320 per m^3 for consumption above 10 m^3 and below 20 m^3, and US$0.356 per m^3 for consumption above 20 m^3. Starting in July 2015, we "linearized" this tariff schedule to avoid having to deal with self-selection of households onto price blocks (as in, e.g., Reiss and White 2005, McRae 2015, and Szabó 2015). With the linearization, every household in the three villages was charged a uniform US$0.236 per m^3 price and the water utility was reimbursed for the revenue loss relative to the block-price schedule. The linearized price structure prevailed throughout

[9] While the baseline survey collected data on 1,488 households, 3 were dropped due to consumption exceeding 50 m^3 for 5 or more months during the experiment; they were most likely operating a business out of their homes. We dropped a further 23 households due to non-participation in the quarterly monitoring surveys (no time-varying demographic information).

the experiment, with randomized discounts applied at various points as described next. In June 2017, however, provincial authorities raised the base tariff to US$0.353 per m^3, a 50 percent increase, which we passed on to all consumers; this change was broadly announced by An Thinh.

The price experiment began in January 2016, 6 months into the price linearization; see Figure 1 for a timeline. All the households surveyed at baseline agreed to participate in the price discount scheme and were randomly divided into three groups. Each group would, according to staggered schedules, receive price discounts of 50 or 75 percent and thus be charged 50 or 25 percent of the July 2015 flat price, respectively (Figure 1).[10] Each price subsidy regime ran for 6 months with the water utility reimbursed for the difference between the official and subsidized prices. However, households were not informed of the overall duration of the price experiment or of any particular discount more than a month in advance.

The experimental design incorporated a six-month pause in all discounts beginning January 2017, one year into the price experiment. The purpose of the pause was to allow a non-parametric test of demand persistence, exploiting the fact that each of the three groups had been exposed to different price discounts over the previous year yet were currently facing the same price. We discuss this test in section 2.4.

Information campaign

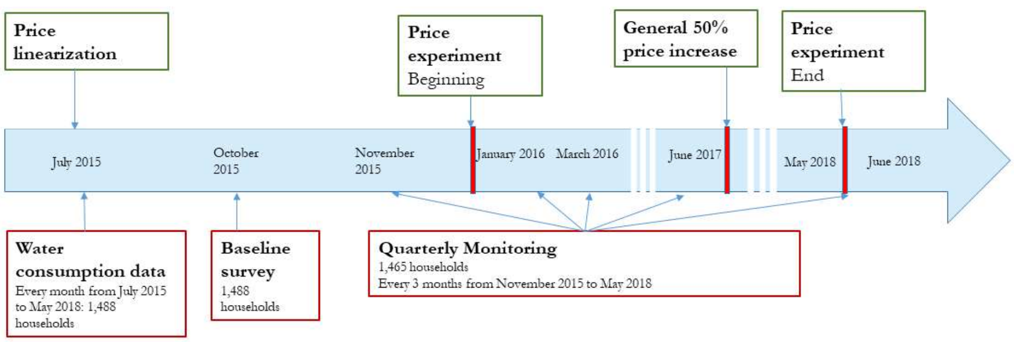
Households were informed of the discounts the month before they took effect in two ways: (1) through a note left on the "Comments" section of the bill, which was handed to each customer upon their monthly in-person payment at a central location in the village, shortly followed by (2) a letter sent by the water utility to each customer's home.[11] To make the discounts particularly salient just before they were to go into effect, any rebate as a result of a discount was applied retroactively to the prior month's bill and consumers were informed that henceforth this would be the prevailing rate.

Price salience was confirmed by households inasmuch as 97 percent of them responded within rounding error of the actual price of US$0.236 when asked in the quar-

[10] The province-wide 50 percent increase in base tariff (June 2017) passed on to the experimental sample is reflected in the lower panel of Figure 1 indicating that a 50 (resp. 75) percent rebate translates into a price equal to 75 (resp. 37.5) percent of the July 2015 price.

[11] Nonpayment or delinquency on water bills is virtually non-existent in our setting.

terly monitoring survey of December 2015 (at the end of the price linearization phase) what price per cubic meter they expected to be charged the following month. Evidently, households were not only well aware of the price of piped water but of the existence of the ongoing linearization scheme as well. Broad awareness of the price flattening is suggested by the fact that even among the 16 percent of households who consumed 10 m^3 or more in December 2015, and who would therefore have paid a higher marginal price absent the linearization scheme, 97 percent still accurately reported within rounding error of the exact price. During this initial phase of the experiment, households were also informed of the imminent subsidy program, although its timing and magnitude were not revealed until the last minute.



Subsidized price rollout

	Jun-Dec 15	Jan-Jun 16	Jul-Dec 16	Jan-Jun 17	Jul-Dec 17	Jan-May 18
Group 1	100%	50%	25%	100%	150%	150%
Group 2	100%	100%	50%	100%	37.5%	150%
Group 3	100%	100%	100%	100%	75%	37.5%

Notes: The table shows the subsidized price rollout, which proceeded as follows: During June-December 2015, all three experimental groups faced 100% of the baseline water price. The discounts began January-June 2016, with Group 1 facing 50% of the baseline price, while Groups 2 and 3 still faced 100%. In the next six month period, Group 1 faced 25%, Group 2 50%, and Group 3 100% of the baseline price, and so on.

Figure 1: Experiment timeline and price subsidy rollout

2.2 Descriptive data

Most of the experimental sample households (70 percent) were connected to piped water in the initial roll-out of An Thinh's network from September 2012 to February

2013, with the remainder connecting throughout the rest of 2013 (10 percent), 2014 (9 percent), and 2015 (11 percent). Before the arrival of piped water, rainwater stored in concrete tanks was the principal source of domestic water supply; for 83 percent of households, internal piping and plumbing fixtures were already in place for conveying rainwater. Appendix Table B.1 shows that all but three households report using piped water at baseline. More than half of the sample (806 households) report still using some rainwater for cooking and drinking (rainwater is filtered for household use and typically boiled before drinking), and half also report using either rainwater or groundwater for shower and bathroom needs. In short, the transition to piped water was far from complete.

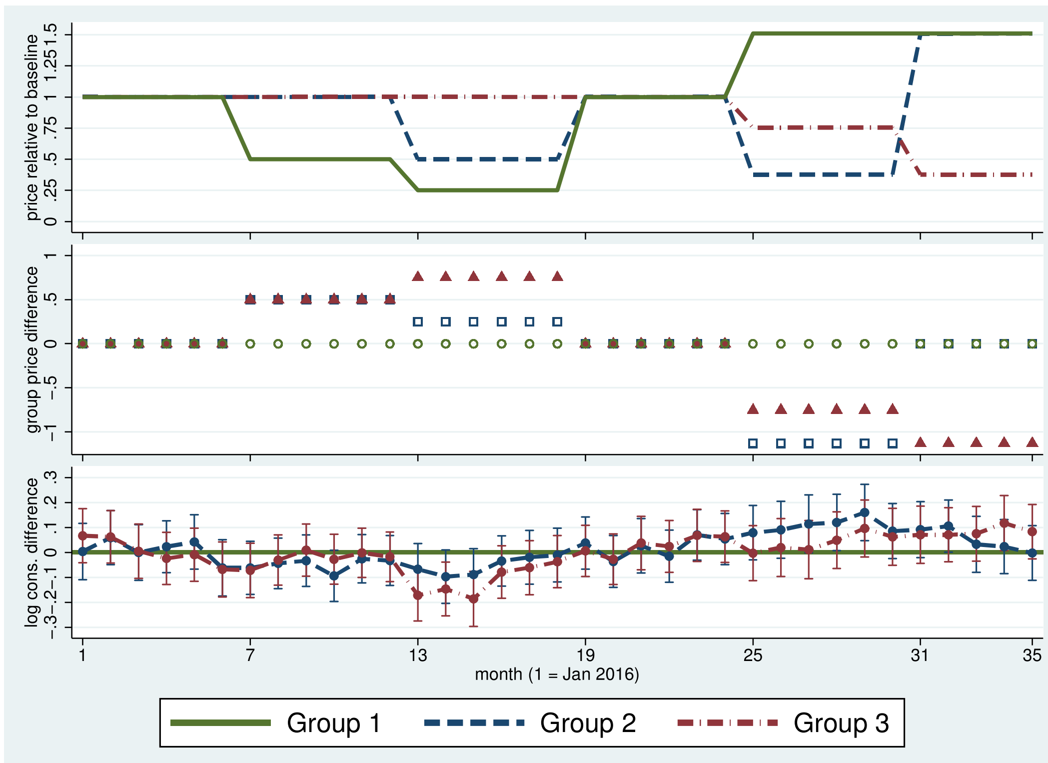
Appendix Table B.2 presents sample descriptive statistics by experimental group as well as balance tests for a range of household characteristics. In no case can we reject balance across experimental groups, including for average monthly water consumption in the pre-subsidy phase (second half of 2015), when all households faced the same water price. Another concern is that, by generating price differences between nearby households and over time, the experiment may lead households to "trade" water among themselves or to store water. There is a strong *prima facie* case against such inter- or intra-household arbitrage in our context. The main benefits of piped water where there is an abundant substitute in the form of rainwater is its higher pressure (for uses such as house cleaning, individual showers and so forth) as well as convenience; i.e., not having to collect, filter, and store rainwater or maintain a water pump, in the case of groundwater. The amenity value of high pressure and convenience would not carry over to water stored or shared between households.[12] That said, in Appendix D, we present statistical tests for arbitrage and find no evidence of the effects of either type on our experimental results.

2.3 Results of price experiment

The upper panel of Figure 2 summarizes the experimental prices faced by each group of households. Recall that during the six months before January 2016 all households

[12]While the option value of not running out of rainwater was also cited as a benefit of piped water, and would be amenable to water sharing, this is an unlikely occurrence for most households.

regardless of experimental group faced the same linear price, which we normalize to 1 in the figure. As also noted, once the experiment was underway, price discounts were rolled out successively to the different groups and then withdrawn, with each discount prevailing for six months. A break in this pattern, which we shall return to shortly, was a six-month pause in all discounts beginning January 2017. Starting July 2017 (month 25), the water utility increased the price by 50 percent, which was fully passed on to the households enrolled in the experiment.



Notes: Top panel: Experimental prices by month and treatment group (base price normalized to 1). Month 1 = July 2015; month 7 = January 2016 (start of experiment); month 19 = January 2017 (start of no discount semester); month 25 = July 2017 (start of general price increase by utility). Middle panel: Group differences in price relative to Group 1. Bottom panel: Treatment effects for groups 2 and 3 (group 1 normalized to zero) by month estimated from a regression of log monthly water consumption on person days, month dummies, and month dummies interacted with the treatment group.

Figure 2: Experimental price and group-wise treatment effects by month

The lower panel of Figure 2 summarizes the experimental results using the base sample of 1,462 households, displaying monthly coefficients b_t and associated 95 percent confidence intervals from the regression

$$\log c_{it} = a_0 + a_1 N_{it} + \sum_k b_t \mathbb{1}_{group=k} + d_t + u_{it}, \tag{1}$$

where $\log c_{it}$ is log monthly water consumption and N_{it} is quarterly person-days present in the household. These monthly treatment effects capture log consumption differences between groups 2 or 3 and the base group 1 (normalized to zero), netting out seasonality and common time shocks and/or trends in water demand using month dummies d_t. For most months of the experiment, the average log consumption of the three experimental groups are statistically indistinguishable, the exception being months 13-15 (group 1 gets the 75 percent discount and thus has higher consumption than control group 3) and months 28-29 (group 2 gets the 75 percent discount and thus has higher consumption than control group 1)

Next, we investigate water customers' response to current month price discounts (see Appendix C for additional specifications). In column 1 of Table 1, we report a regression of (log) monthly water consumption on dummies for the two discounts, household person-days (quarterly), month dummies, and household fixed effects. There is a small, but precisely estimated, increase in water consumption in response to the discounts, significantly larger for the 75 percent discount than for the 50 percent discount, but still under 10 percent. In column 5, we estimate a static (semi-log) demand curve using the current month's piped water price, yielding a precisely estimated price elasticity of 0.076.[13]

2.4 Persistence in piped water demand

Reduced-form evidence

Having established an overall negative response of monthly piped water consumption to the contemporaneous price across the 35 months of the experiment, we now test for demand persistence. We do so by asking whether conditional on the contemporaneous price, current consumption is affected by the past price (and thereby by past consump-

[13] Aside from the restriction that a semi-log demand curve imposes on the discount dummy coefficients, the price specification also uses variation induced by the general tariff increase of July 2017.

Table 1: Price discounts and piped water consumption

	Dependent variable: log monthly piped water consumption							
		In preceding m months				In preceding m months		
		$m=1$	$m=2$	$m=3$		$m=1$	$m=2$	$m=3$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
50% discount (current mo.)	0.034	0.027	0.028	0.026				
	(0.015)	(0.018)	(0.018)	(0.017)				
	[0.021]	[0.135]	[0.111]	[0.132]				
75% discount (current mo.)	0.085	0.067	0.057	0.059				
	(0.015)	(0.019)	(0.018)	(0.017)				
	[0.000]	[0.000]	[0.001]	[0.001]				
50% discount (past mo.)		0.008	0.007	0.010				
		(0.017)	(0.015)	(0.014)				
		[0.646]	[0.629]	[0.482]				
75% discount (past mo.)		0.022	0.035	0.030				
		(0.017)	(0.015)	(0.015)				
		[0.212]	[0.022]	[0.042]				
Current water price					-0.076	-0.049	-0.040	-0.048
					(0.015)	(0.019)	(0.018)	(0.017)
					[0.000]	[0.010]	[0.026]	[0.006]
Past water price (minimum)						-0.031	-0.046	-0.038
						(0.017)	(0.017)	(0.017)
						[0.073]	[0.006]	[0.026]
Observations	46,784	45,322	43,860	42,398	46,784	45,322	43,860	42,398
Clusters	1,462	1,462	1,462	1,462	1,462	1,462	1,462	1,462
R^2	0.659	0.661	0.663	0.666	0.659	0.661	0.663	0.666
Month dummies	YES	YES	YES	YES	YES	YES	YES	YES
HH fixed effects	YES	YES	YES	YES	YES	YES	YES	YES

Notes: Robust standard errors in parentheses clustered at household level; p-values in square brackets. Columns 1-4 include dummies for whether the household is currently receiving the 50 percent or 75 percent discount as well as (in cols. 2-4) whether the household received 50 percent or 75 percent discount in the previous month(s). Columns 6-8 include the minimum water price faced over the past 1, 2, and 3 months, respectively. All regressions control for the number of person-days.

tion). Thus, we extend the specifications reported in columns 1 and 5 of Table 1 by adding dummies for whether the household ever received each type of discount in the preceding 1, 2, or 3 months (resp. cols. 2-4). For the latter two windows, we find that the effect of a past 75% discount is statistically significant at above the 5% level, whereas the effect of a past 50% discount is insignificant for all time windows. For an analogous exercise using the semi-log demand specification of column 5, we add the minimum price prevailing over the past 1, 2, or 3 months (resp. cols. 6-8). Here, the null hypothesis of zero effect of past water prices conditional on the current water price can be rejected with p-values of, respectively, 0.073, 0.006, and 0.026. While the percentage effects of present and past price discounts on monthly water consumption are all rather small, as are the corresponding price elasticities, we find statistical evidence for demand persistence.

Our experimental design also affords a simple test of demand persistence that can be read directly from the bottom panel of Figure 2. First, as seen in the top panel, during months 7-18 each of the groups received a different sequence of price discounts to induce different average consumption levels by month 18. Given persistence, these exogenous differences in past consumption should induce corresponding differences in month 19 consumption, during which month no group received a price discount. Note that this test is conceptually identical to the regression-based tests just presented in Table 1 but focuses on one month of current consumption data. As it happens, the bottom panel of Figure 2 shows no detectable difference in consumption in month 19 between group 1, the most subsidized of the three during months 7-18, and group 3, the least subsidized. We attribute this non-result to a lack of power.[14]

Mechanisms

Both provincial water authorities and the director of An Thinh utility told us they expected households to increase piped water consumption over time. Indeed, in its first 3 years of operation (preceding our experiment), piped water use among An Thinh's customers was growing by an average of 0.9 percent per month.[15] While households had previously collected rainwater that needed to be stored in and pumped from dedicated tanks, piped water is more convenient and provides higher pressure; it also does not need to be filtered and boiled for cooking purposes like rainwater. As these benefits become manifest through continued use, An Thinh's customers were expected to increasingly shift out of rainwater and into piped water, especially for intensive uses like showers, cleaning, and so forth.[16] Whether we call this phenomenon habit formation or

[14] Our test was powered using the plausible value of 0.8 for the static price elasticity of demand for piped water (see Diakité et al. 2009 and Appendix E). We have just seen, however, that the price elasticity is only about a tenth of this magnitude, rendering our simple test for persistence substantially under-powered ex-post. Because of the dissipation of demand persistence over time, power is substantially lower for the same test performed using month 20 consumption (see Appendix E), and even lower using month 21, and so on.

[15] This is a within household estimate and thus accounts for selectivity into An Thinh's customer-base by connection month.

[16] Monthly water consumption during the hot season from May to September averages about 12 percent higher than during the rest of the year in our sample, largely reflecting the greater demand for showers.

learning through experience, or some combination thereof,[17] it implies that exogenously inducing greater piped water consumption in the past should raise today's consumption conditional on the current price, which is precisely what we find in Table 1.

Alternative explanations for this same evidence include: (i) lagged awareness of price discounts; (ii) loss aversion; and (iii) complementary investment in piped water infrastructure, none of which are likely to be quantitatively important in our setting, as we now discuss. Price awareness or salience was already mentioned in the context of our experimental design. Water customers were informed of the discounts just in advance in two attention-getting ways, on their water invoice handed to them during in-person bill-paying and in an official letter sent to their homes. Through this publicity campaign, we minimized the lack of awareness or inattention to the discounts as they came into force.

If loss-averse households form a mental account and/or a reference point for piped water expenditures, then the impact of the current price would, in general, depend on the past price; in particular, on whether the former is higher or lower than the latter (see, e.g., Thaler 1985; Kőszegi and Rabin 2006). While such a phenomenon could account for the evidence of demand persistence in Table 1, loss aversion also has the directly testable prediction that demand responses are more elastic to price increases than to price decreases (see, e.g., Ahrens et al. 2017). In Appendix F, however, we present evidence to the contrary, finding that price increases and decreases have symmetric effects on piped water use. Thus, loss aversion is not relevant in our setting and, hence, cannot be an important source of demand persistence. More broadly, unlike the habit formation or experience good mechanism, loss aversion cannot explain rising water consumption in a static price environment, as prevailed before our experiment.

Investment in piped water infrastructure is also unlikely to be driving the results in Table 1. First, we find persistence over as little as 2 months, which would not be

[17] Consumers may also be learning about the *reliability* of the utility's service over time through experience and, insofar as their initial presumption is that reliability is poor, might increase consumption as a consequence. We think it unlikely, however, that this is driving demand persistence throughout our experiment. Most households had already been connected for a couple of years by that time and thus were quite familiar (and satisfied) with the water utility's performance. Indeed, from our quarterly monitoring survey administered in December 2015, just before the start of the price experiment, we find that only about one percent of households had complaints about the reliability of the service (interruptions and/or lack of water pressure, to use the most expansive definition of "reliability") over the preceding 3 months.

enough time for households benefiting from a sudden price discount to install plumbing inside their homes or build new showers. Second, as noted previously, most households already had such infrastructure installed before their connection to the piped water service. Third, and perhaps most telling, the absolute magnitude of the price discounts and their uncertain duration, while sufficient to marginally incentivize monthly piped water consumption in the treated group as compared to the control, are unlikely to induce differential big-ticket investments or appliance purchases (e.g., a washing machine) across these groups, especially when the average monthly water bill, at the undiscounted price, is less than US$2.[18]

3 Endogenous Preferences for Piped Water

In light of our experimental findings, we next specify a model of consumer demand for piped water that allows for preferences to change with past consumption and yet is empirically tractable as well as convenient for our investigation of optimal pricing. With such a model, we will able to move beyond the reduced-form persistence parameters identified earlier and estimate preference parameters, which will allow subsequent welfare analyses.

3.1 Setting the stage

We begin with quasi-linear instantaneous utility with time-varying preference parameter ρ_t

$$u_t(c, z) = v(c; \rho_t) + \alpha z, \tag{2}$$

where c and z are current piped water and numeraire consumption, respectively, and sub-utility function v has properties $v' > 0$ and $v'' < 0$. For convenience, we choose v of the form

$$v(c; \rho_t) = \rho_t c - c \log c \tag{3}$$

[18]Two additional pieces of evidence militate against the investment channel: (1) As already noted, we cannot reject symmetry of responses to increases and decreases in price (Appendix F). In a durable investment model, however, reductions in price should lead to *larger* demand responses than increases in price insofar as reversibility of investment is costly; (2) The investment mechanism also implies a greater response to price discounts among wealthier households, who can more readily afford appliance purchase, but we cannot reject equal price responses across wealth groups (Appendix C.2).

to deliver semi-log demand.[19]

The parameter ρ_t has a time-invariant component θ, an exogenous time-varying and mean-zero component ξ_t, and an endogenous time-varying component $\bar{C}_{t-1}$ as follows:

$$\rho_t = 1 + \theta + \beta \bar{C}_{t-1} + \xi_t, \tag{4}$$

where we assume that $\beta \geq 0$. A nonzero β can be rationalized by a model of habit formation (Pollak, 1970; Becker and Murphy, 1988) or of experience goods (Nelson, 1970; Shapiro, 1983).

For the sake of tractability, we assume that $\bar{C}_t$ follow law of motion

$$\bar{C}_{t-1} = \sum_{k=1}^{\tau} \gamma^{k-1} \log c_{t-k}, \tag{5}$$

where τ is a finite number of periods and $\gamma \in (0, 1)$. In other words, consumption in period t is assumed to directly affect future preferences for τ periods onward. A geometric lag, moreover, implies that the sensitivity of future preferences to past consumption decays at a constant rate over time; when $\tau = 1$, we simply have $\bar{C}_{t-1} = \log c_{t-1}$.

3.2 Water consumption choices

A "rational addict" (Becker and Murphy, 1988) or a "strategic experimenter" (Bergemann and Välimäki, 2006) internalizes the effect of current consumption on both current and future preferences. Optimal consumption choice, in particular, takes prices $\{p_t\}$ as given and solves dynamic program

$$V_t(\theta, \bar{C}_{t-1}, p_t, y_t | \xi_t) = \max_{\{c_{t+k}, z_{t+k}\}_{k \geq 1}} \mathbb{E}_t \sum_{k \geq 0} \delta^k u(c_{t+k}, z_{t+k} | \theta, \bar{C}_{t-1+k}, \xi_{t+k}) \tag{6}$$

subject to the law of motion given by (5) and the budget constraint, $\forall k \geq 1$,

$$z_{t+k} + p_{t+k} \cdot c_{t+k} = y,$$

[19]Similarly, Alvarez and Argente (2020) assume quasi-linear preferences and a convenient form for v to estimate consumer surplus for one good, which, in their setting, is Uber rides.

where y is consumer income, which we assume to be time-invariant. Recognizing, however, that consumers might not be fully sophisticated about how their future preferences are shaped by current choices, we posit instead that consumers are naive or myopic, i.e., unaware of process described by equation (5), and thus behave to solve the static program

$$V_t(\theta, \bar{C}_{t-1}, p_t, y|\xi_t) = \max_{c,z} u(c, z|\theta, \bar{C}_{t-1}, \xi_t) \quad (7)$$

subject to budget constraint

$$z + p_t \cdot c = y.$$

Our assumption of static optimization is an application of projection bias (Loewenstein et al., 2003) whereby agents perceive that their future preferences will be identical to their current preferences. Whether one interprets our preference structure in terms of habit formation or experience good consumption, projection bias equally predicts that consumers underestimate their future taste for piped water; either they are oblivious to their future habits or they are pessimistic about how much they will value the service after having experienced it. Loewenstein et al. (2003) cite a range of research providing suggestive evidence of projection bias. More recent experimental studies also provide powerful evidence against substantial sophistication (Augenblick and Rabin, 2019; Acland and Levy, 2015).[20] In addition, and more generally, since our bounded-rationality assumption amounts to a functional form restriction on the intertemporal Euler equation, we can provide a formal statistical test of naive/myopic/static versus fully sophisticated/forward-looking/dynamic optimization in our setting. The result of this test, reported in Appendix H, indicates that piped water consumers do not appreciably internalize the future effects of their current consumption, whether due to habit formation or to learning.

With a binding budget constraint, optimal consumption is given by first-order condition:

$$\log c_t(\theta, \bar{C}_{t-1}, p_t|\xi_t) = \theta - \alpha p_t + \beta \bar{C}_{t-1} + \xi_t. \quad (8)$$

[20]In a carefully designed experiment, Hussam et al. (2017) find a degree of rational addiction to handwashing in India. As they note, however, "our design sets up the optimal scenario to facilitate rational habit formation: households are fully aware that we want to help them develop a habit of handwashing, and we reiterate the future dates at which the value of the behavior will change." It is unclear, therefore, how this sophistication result generalizes to a setting like ours in which households are not primed for habit change.

Equation (8) defines the short-run demand function as in Pollak (1970). The short-run price elasticity (actually semi-elasticity, or elasticity at $p_t = 1$) is given by α, whereas β captures the extent to which habits formed in the past influence today's consumption, which we thus assume to be nonnegative.

3.3 Long-run demand

While consumer short-run demand is given by equation (8), we define long-run consumption by first adding time $t = 0$, a pre-connection period, and then write

$$\log c_\infty(\theta, p) = \lim_{t\to\infty} \mathbb{E}_0 \log c_t(\theta, \bar{C}_{t-1}, p|\xi_t), \tag{9}$$

i.e. the expectation is taken before connection at $t = 0$ of the limit value of (log) consumption, given that price p is time-invariant. Further, given that $\beta \geq 0$ and $\gamma \in (0,1)$, we have $\beta\frac{1-\gamma^\tau}{1-\gamma} \leq 1$, so that

$$\lambda \equiv \left[1 - \beta\frac{1-\gamma^\tau}{1-\gamma}\right]^{-1} \geq 1. \tag{10}$$

We now have (proof in Appendix):

Lemma 1 If inequality (10) holds, then long-run demand given by (9) is well-defined and

$$\log c_\infty(\theta, p) = \lambda(\theta - \alpha p). \blacksquare \tag{11}$$

As the stock of consumption builds, demand increases but so does its price elasticity (from α to $\lambda\alpha$). Intuitively, a long-run increase in price not only decreases contemporaneous consumption but also reduces past consumption, thereby amplifying the static price response.

4 Structural estimation

In this section, we take up the estimation of λ and the other preference parameters using data from the experiment and then validate the structural model on pre-experimental data.

4.1 Preference Estimation

Empirical specification

Equation (8) describes consumer behavior; its empirical counterpart is

$$\log c_{it} = -\alpha p_{it} + \beta \sum_{k=1}^{\tau} \gamma^{k-1} \log c_{it-k} + \omega N_{it} + \theta_i + \zeta_t + \varepsilon_{it}, \tag{12}$$

where i indexes consumers. We decompose the mean-zero preference shock $\xi_{it} = \omega N_{it} + \zeta_t + \varepsilon_{it}$ into an observed component depending on the (time-varying) deviation from long-run household size N_{it}, an aggregate shock ζ_t, and an unobserved component ε_{it}, with all these components having mean zero. Note that price p_{it} varies over time and (randomly) across consumers by experimental group.

To remove the consumer fixed effect θ_i, we difference (12) between period t and $t-L$, for some to-be-determined time lag L, such that

$$\Delta^L \log c_{it} = -\alpha \Delta^L p_{it} + \beta \sum_{k=1}^{\tau} \gamma^{k-1} \Delta^L \log c_{it-k} + \omega \Delta^L N_{it} + \Delta^L \zeta_t + \Delta^L \varepsilon_{it}, \tag{13}$$

where, e.g., $\Delta^L p_{it} \equiv p_{it} - p_{it-L}$ and $\Delta^L \zeta_t$ is a month fixed effect (as included in the reduced form models above). Any choice of lag used for differencing is equally valid econometrically. Below, we use a data-driven approach to select L.

Identification

Experimentally induced variation in the price of piped water can be used to identify, not only the short-run demand curve, but also the causal impact of past water consumption on current demand. Consider identification of the key preference parameters α and β in the case $\tau = 1$; the argument is the same when $\tau > 1$. Equation (13) becomes

$$\Delta^L \log c_{it} = -\alpha \Delta^L p_{it} + \beta \Delta^L \log c_{it-1} + \omega \Delta^L N_{it} + \Delta^L \zeta_t + \Delta^L \varepsilon_{it}, \tag{14}$$

from which we see that, *conditional* on the first lagged consumption change, the change in current consumption depends only on the current price change $\Delta^L p_{it}$ (which identifies α) but not on the change in lagged prices $\Delta^L p_{it-k}$ or any function of lagged prices for $k \geq 1$. This is the potential set of theoretically valid exclusion restrictions. Theory, moreover, implies that the lagged consumption change $\Delta^L \log c_{it-1}$ (unconditionally on $\Delta^L \log c_{it-2}$) *does* depend on the change in lagged prices or functions thereof. Hence, the intertemporal dependence parameter β is identified in principle.

The general case

When $\tau \geq 1$, we want to select an optimal τ and to recover γ along with the other parameters of equation (13). Conditional on values of γ and τ, the parameters α, β, and ω are estimable and identified as already discussed. Given such estimates, we may compute residuals $\hat{e}_{it} = \Delta^L \hat{\varepsilon}_{it}$ and form the concentrated sum of squared errors

$$S(\gamma, \tau) = \sum_{i=1}^{N} \sum_{t=1}^{T} \hat{e}_{it}^2(\gamma, \tau). \tag{15}$$

We minimize $S(\gamma, \tau)$ by iterating between $\tau = 1, 2, 3, 4$ and a grid search over γ (in increments of 0.01) conditional on τ, the latter step equivalent to nonlinear least squares.[21]

Our estimation consists of the following sequence of steps, beginning with data-driven algorithms for choosing both lag-length L for differencing and the lagged price instruments:

1. With $\tau = 1$ and using a 'basic' instrument set $\{\Delta^L p_{it-1}, \Delta^L \log p_{it-1}\}$ for $\Delta^L \log c_{it-1}$, select the L that yields the best-fitting first-stage for equation (13). Call this L^*.

2. Given L^*, use machine-learning (post-double selection methodology of Belloni et al., 2012) to select a sparse instrument set from a successively expanded basic instrument set, i.e., $\{\Delta^{L^*} p_{it-1}, ..., \Delta^{L^*} p_{it-k}, \Delta^{L^*} \log p_{it-1}, ..., \Delta^{L^*} \log p_{it-k}\}$.

3. Repeat step 1 with sparse instrument set replacing basic set to check the robustness of L^*.

[21] Such nested minimization, also known as concentration or profiling, is a standard computational shortcut for certain nonlinear regression models as discussed in Hansen (2022).

4. Given L^* and the sparse instrument set, simultaneously choose τ^* (restricting $\tau < 5$ as a practical matter) and estimate γ by minimizing (15).

Inference

Inference on the structural parameters must account for the uncertainty induced by the selection of lag-length, instruments, and lag order in steps 1-4. We thus use a nonparametric bootstrap of our entire estimation procedure to construct confidence intervals and to perform hypothesis testing. One potential threat to the asymptotic validity of such a bootstrap, however, is the non-identification of γ under the null $\beta = 0$.[22] We thus proceed as follows: First, we set $\gamma = 0$ and perform the asymptotically valid test for $\beta = 0$, i.e., by bootstrapping the entire estimation procedure up to step 3 (this test should have reasonable power so long as the true γ is not too far from 0). If we reject the null, then we conclude that the full model is identified and hence the nonparametric bootstrap of our full procedure is asymptotically valid.[23,24]

Steps 1-3, documented in Appendix G, yield $L^* = 8$ and a sparse instrument set that coincides with the basic set, as well as the structural parameter estimates reported in column 1 of Table 2 (i.e., for the $\tau = 1$ case). The null hypothesis $\beta = 0$ is rejected with a bootstrap p-value$= 0.005$, justifying our use of the bootstrap for inference in the more general model.

4.2 Estimating γ and the long-run demand elasticity

Turning now to estimation step 4, we obtain $\tau^* = 3$ and the parameter estimates reported in column 2 of Table 2. The bias-corrected bootstrap confidence intervals for the key structural parameters (α, β, γ) all lie above zero. Thus, we find that greater piped water consumption over the last three months encourages current consumption, although with $\hat{\gamma} = 0.41$ the influence of consumption three months ago is only about

[22] Neither the multiplier bootstrap suggested by Hansen (1996) nor the asymptotic approach of Andrews and Cheng (2012) can account for model selection uncertainty.

[23] More precisely, conditional on the choices of L^*, instruments, and τ^*, our nonlinear least squares or M-estimator yields asymptotically normal estimates under the usual regularity conditions and, consequently, the nonparametric bootstrap is consistent (Horowitz, 1997). The validity of the bootstrap in dealing with model selection uncertainty has been shown by, e.g., Kilian (1998).

[24] This sequential testing procedure is pointwise valid for data-generating processes with β far away from zero (see Andrews and Cheng 2012 for a uniformly valid procedure).

Table 2: Preference Estimation

		$\tau = 1$	$\tau = 3$
α	(SR elasticity)	0.042	0.046
		[0.019,0.077]	[0.022,0.074]
β		0.616	0.402
		[0.230,0.896]	[0.266,0.604]
γ		—	0.41
			[0.19,0.610]
ω		0.601	0.633
		[0.058,1.035]	[0.025,1.049]
$\alpha\lambda$	(LR elasticity)	0.109	0.126
		[0.054,0.217]	[0.065,0.508]
$H_0 : \alpha = \alpha\lambda$		0.054	0.103
Bootstrap p-value			
Month dummies		YES	YES
Observations		35,087	33,522
Clusters		1,462	1,462

Notes: Cluster bootstrapped bias-corrected percentile 95% confidence intervals in square brackets. Dependent variable is $\Delta^{L^*} \log c_{it}$, where $L^* = 8$. Estimation conditional on γ in column (2) is by two-step GMM with (excluded) instruments Δp_{t-1} and $\Delta \log p_{t-1}$ and estimation of γ is by nonlinear least squares.

($0.41^2 =$) one-sixth as great as that of previous month's consumption. Based on the column 2 estimates, the short-run price elasticity (at the baseline price of 1) is $\hat{\alpha} = 0.046$, the persistence parameter is $\hat{\lambda} = \left[1 - \hat{\beta}\frac{1-\hat{\gamma}^{\tau}}{1-\hat{\gamma}}\right]^{-1} = 2.74$ and, consequently, the long-run elasticity (lemma 1) is $\hat{\alpha}\hat{\lambda} = 0.126$, as compared to 0.109 in column 1 where $\tau = 1$. A one-sided test of the null hypothesis of equal short- and long-run elasticities in column 2 has a bootstrap p-value of 0.103, as compared to 0.054 in column 1, where the long-run elasticity calculation involves fewer estimated parameters.

Intuition for the identification of demand persistence is most easily seen for $\tau = 1$, in which case $\lambda = 1/(1 - \beta)$. Consider "reduced-form" regression

$$\Delta^L \log c_{it} = b_1 \Delta^L p_{it} + b_2 \Delta^L p_{it-1} + u_{it}, \tag{16}$$

with $\Delta^L N_{it}$ and month dummies included but omitted from (16) for expositional convenience. We find that $\hat{b}_1 = -0.050$ and $\hat{b}_2 = -0.042$.[25] Although both of these

[25] The robust (clustered) asymptotic standard errors are, respectively, 0.024 and 0.022.

reduced-form price elasticities are small, like those in columns 6-8 of Table 1, their ratio $\hat{b}_2/\hat{b}_1 = 0.839$ is large. This ratio is an upper bound on β, our estimate of which is 0.616 (Table 2, col. 1). To see the connection, note that equation (14) implies (again suppressing terms involving $\Delta^L N_{it}$ and month dummies) that

$$\Delta^L \log c_{it} = -\alpha\Delta^L p_{it} - \beta\alpha\Delta^L p_{it-1} + \beta^2\Delta^L \log c_{it-2} + \Delta^L \varepsilon_{it}. \quad (17)$$

Thus, regression (16) evidently omits $\Delta^L \log c_{it-2}$ from the true structural model (equation 17). Since $cov(\Delta^L p_{it-1}, \Delta^L p_{it-2}) > 0$ and $cov(\Delta^L \log c_{it-2}, \Delta^L p_{it-2}) < 0$, $|\hat{b}_2|$ overestimates $|\beta\alpha|$ and, therefore, $\hat{b}_2/\hat{b}_1$ overestimates β.[26]

To summarize, exploiting experimental price variation, we find evidence of substantial intertemporal dependence in piped water consumption. The relatively large estimate of the persistence parameter λ implies that small contemporaneous price responses are greatly magnified over time with potentially important welfare implications. Indeed, as we will show in Section 5, $\lambda \gg 1$ implies significant welfare gains to maintaining low marginal prices and to deferring lump-sum payment until long-run preferences are formed.

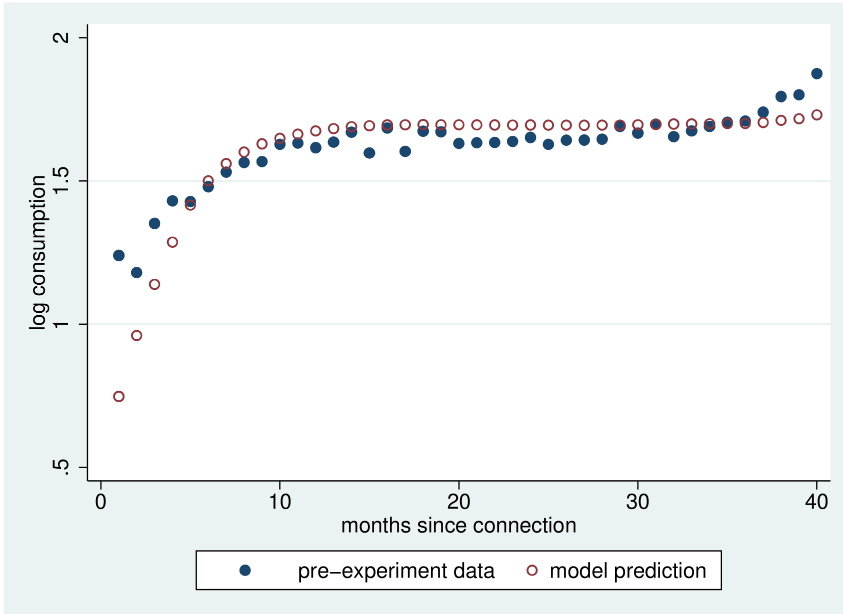
4.3 Model validation: out-of-sample consumption dynamics

Next, we assess how well the structural model captures piped water consumption dynamics out-of-sample. Connections to the new water utility in My Huong commune began in September 2012 and continued up until our pricing experiment started in January 2016 (see Appendix Figure B.1). We use our estimates $\hat{\alpha}, \hat{\beta}, \hat{\gamma}$, and $\hat{\theta}_i$ to predict water consumption for each household from the first month of their connection ($t = 1$) up until December 2015, assuming that $\bar{C}_0 = 0$ and computing piped water demand under the prevailing official block tariff schedule.[27]

In Figure 3, the dots show monthly predicted margins from a regression of *actual* log monthly consumption on month since connection dummies, which also controls for

[26]Had we randomized the discounts in our experiment *every month*, so that $cov(\Delta^L p_{it-1}, \Delta^L p_{it-2}) = 0$, then we could have recovered β directly from the ratio of the reduced form price elasticities $\hat{b}_2$ and $\hat{b}_1$.

[27]While time-varying demographics, N_{it}, were not collected before the experiment, the estimated household fixed effect $\hat{\theta}_i$ incorporates the household average demographics $\bar{N}_i$ throughout the experiment.



Notes: Solid dots represent average log piped water consumption by month since connection, partialling out (estimated) preferences and month × year fixed effects, using data from Sept. 2012–Dec. 2015. Hollow dots represent model-based predictions.

Figure 3: Actual vs. model-generated pre-experiment log consumption

preferences $\hat{\theta}_i$ and for calendar month × year fixed effects to absorb aggregate shocks.[28] Water usage data cover the first 40 months of the utility's operation in the commune prior to our experiment and includes all (and only) the 1462 households in the experimental sample. The dashed curve in Figure 3, derived from exactly the same regression procedure but with the dependent variable now being model-generated log monthly consumption, broadly fits the time pattern in the actual data. Most encouraging is that model and data seem to agree on how long it takes, on average, for consumption to reach its steady state level starting from the initial month of connection; a static model of demand, by contrast, would trivially, but counter-factually, predict that steady state consumption is achieved at $t = 1$. The reasonably good match between pre-experiment data and model also suggests that alternative mechanisms potentially leading to growth in piped water consumption but unlikely, for reasons already discussed, to have generated persistence *within* the experiment (e.g., appliance investment, changing perceptions of service reliability) were similarly unimportant *prior* to the experiment.

[28]Note that, in an alternative household fixed specification, we could not distinguish cohort (month of connection) and time (aggregate shock) effects.

5 Optimal pricing of a new utility service

We are now ready to investigate the optimal pricing of a new utility service. Returning to the model of Section 3, the *ex-ante* expected indirect utility, i.e., prior to consumption of the service, may be defined as

$$\mathbb{V}_0(\theta, p, y) = \frac{1-\delta}{\delta}\left[\alpha y + \mathbb{E}_0 \sum_{t\geq 1} \delta^t V_t(\theta, \bar{C}_0, p, y|\xi_t)\right]. \tag{18}$$

Ex-post utility, i.e., once consumption has reached its long-run steady state level, is similarly given by

$$\mathbb{V}_\infty(\theta, p, y) = \frac{1-\delta}{\delta}\left[\alpha y + \mathbb{E}_0 \sum_{t\geq 1} \delta^t V_t(\theta, \bar{C}_\infty, p, y|\xi_t)\right]. \tag{19}$$

In defining these utilities, we abstract from transition dynamics. Thus, equation (18) assumes that $\bar{C}_0 = 0$ and remains so over time, whereas equation (19) assumes that $\bar{C}_{t-1}$ reaches $\bar{C}_\infty$ in period $t = 1$. Assuming further that the preference shock ξ_t is "small", we obtain:[29]

$$\mathbb{V}_0(\theta, p, y) \approx e^{\theta - \alpha p} + \frac{1}{\delta}\alpha y \tag{20}$$

and

$$\mathbb{V}_\infty(\theta, p, y) \approx e^{\lambda(\theta - \alpha p)} + \frac{1}{\delta}\alpha y. \tag{21}$$

Before turning to optimal pricing decisions, we specify the utility's cost structure. Let K denote per customer fixed cost, which includes plant and equipment, operating costs, and home connection costs. To avoid dependence on discount rate δ, we interpret K as a perpetual flow starting once the utility begins operating in period 1. There is also a marginal cost c for variable inputs such as fuel for pumping and chemicals for treating the water. The optimization program for the public utility is thus to maximize aggregate welfare under the constraint that fees and water sales revenues per customer cover the utility's average cost.

[29]Note that for every $t \geq 1$, $V_t(\theta, 0, p, y|\xi_t) = e^{\theta - \alpha p + \xi_t} \approx (1 + \xi_t)e^{\theta - \alpha p}$. Thus, $\mathbb{E}_0 V_t(\theta, 0, p, y|\xi_t) \approx e^{\theta - \alpha p}$, which is time-invariant and hence yields expression (20). A similar argument holds for (21).

5.1 Optimal pricing: homogeneous preferences

To isolate the implications of demand persistence for optimal pricing, we first consider the homogeneous preference case, in which the entire population shares the same value of θ. Before setting up the water utility to serve this population, we suppose that the social planner undertakes a preference elicitation to determine both θ and α. The "unsophisticated" planner, however, does not anticipate demand persistence, assuming rather that $\lambda = 1$.

To make our analysis of optimal pricing independent of discount rate δ and consistent with our definition of average cost K, we express the lump-sum connection fee as a flow of payments F starting in period 1, which is the same as a flow of payments δF starting in period 0.[30] In equating the upfront payment of the fee to perpetual installment payments of δF on a loan of size F taken at (monthly) interest rate $\frac{1}{\delta}$, we abstract from credit constraints. While credit constraints may limit what utilities can charge upfront for a connection (Devoto et al., 2012a; Lee et al., 2020; Berkouwer and Dean, 2021), the relevant policy instrument to deal with this issue is not the utility's price structure but rather the provision of loans to households.

A generalization of the two-part tariff

Consider, first, the unsophisticated planner, who does not account for endogenous preferences. As is well-known, the utility's markup over marginal cost c in this scenario will be zero, so that $p^* = c$, and the utility will cover its average costs, insofar as it can, entirely through a lump-sum connection fee.[31] Ex-ante willingness-to-pay of consumers is, therefore, given by $\mathbb{V}_0(\theta, c, 0) = e^{\theta - \alpha c}$. It follows that a decision to undertake the project based on ex-ante willingness-to-pay will only hold for fixed cost $\alpha K \leq e^{\theta - \alpha c}$, which brings surplus $\mathbb{V}_\infty(\theta, c, 0) = e^{\lambda(\theta - \alpha c)}$. This case is represented by the dotted line up to point A in Figure 4, showing consumer surplus at different levels of fixed cost

[30] To see why, note that a one-time connection fee F assessed at time $t = 0$ is equivalent to a flow of payments equal to $(1 - \delta)F$ paid in every period or to a flow of payments equal to $\frac{1-\delta}{\delta}F$ paid from $t = 1$ onward. By the same token, a flow of payments equal to F paid in every period starting at $t = 1$ is equivalent to payments equal to δF made from $t = 0$ onward.

[31] An optimal tariff schedule does not necessarily have only two parts. However, if the demand elasticity is constant across consumers, which is the case in our setting (as we establish in Appendix C), then the optimal nonlinear price is a two-part tariff (provided that the social planner does not have a redistribution motive).

αK.

Now consider a sophisticated social planner who correctly anticipates future preferences and thus maximizes $\mathbb{V}_\infty(\theta, p, -\delta F)$, or

$$\max_{\{F,p\}} e^{\lambda(\theta-\alpha p)} - \alpha F, \tag{22}$$

subject to the consumer participation constraint reflecting ex-ante preferences

$$\mathbb{V}_0(\theta, p, -\delta F) = e^{\theta-\alpha p} - F \geq 0 \tag{23}$$

and the utility's budget constraint

$$(p-c)e^{\lambda(\theta-\alpha p)} + F \geq K. \tag{24}$$

Revenues, on the left-hand side of equation (24), are based on long-run consumption, which the social planner correctly forecasts so that in this respect both maximand (22) and the budget constraint differ from those of an unsophisticated social planner.

With a binding budget constraint, the planner's problem reduces to

$$\max_p [1 + \alpha(p-c)]e^{\lambda(\theta-\alpha p)} - \alpha K. \tag{25}$$

Assuming a non-binding participation constraint, the first-order condition implies that

$$p^* - c = \frac{1}{\alpha}\left(\frac{1}{\lambda} - 1\right), \tag{26}$$

indicating a negative markup for $\lambda > 1$. Intuitively, a marginal reduction in price increases long-run consumer surplus, amplified (as the planner correctly anticipates) by the persistence parameter λ, but also increases the lump-sum payment F required to keep the budget in balance. The conventional two-part tariff emerges as a special case with $\lambda = 1$; in the absence of long-run surplus amplification, a sub-marginal cost price is inefficient.

A sophisticated social planner's negative markup thus increases consumer surplus (e.g., at $\alpha K = e^{\theta-\alpha c}$ from point A to point B) and expands the range of project viability. Price p^* and connection fee F^* prevail until the consumer's participation constraint binds (point C), i.e. $e^{\theta-\alpha p^*} = F^*$. The budget constraint can then be written:

$$(p-c)e^{\lambda(\theta-\alpha p)} + e^{\theta-\alpha p} \geq K \tag{27}$$

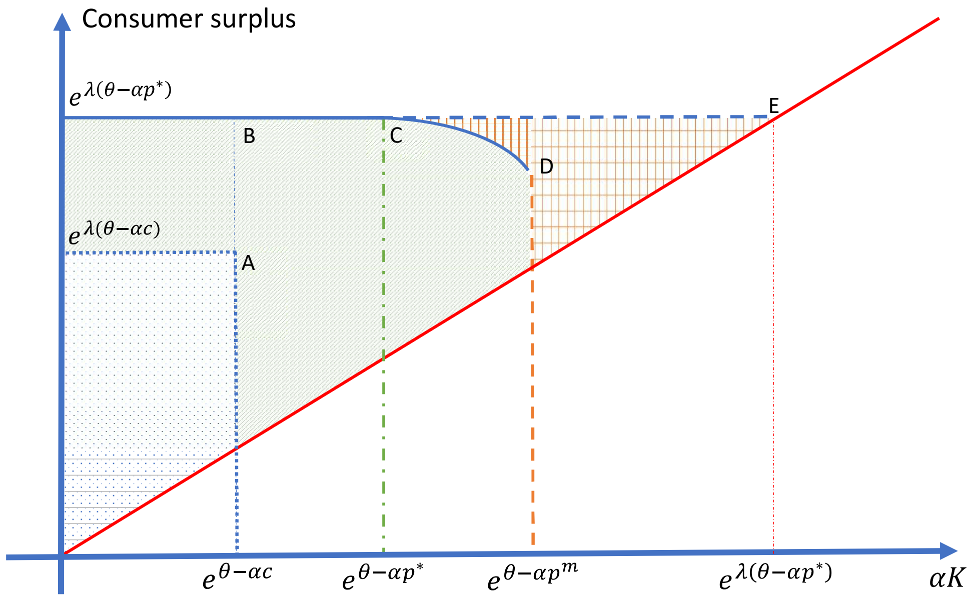
The left-hand side of inequality (27) is the utility's revenue function when the consumer's participation constraint is binding. When this function is increasing at $p = p^*$, the utility can raise revenues by increasing price and in doing so lower F to induce connection.[32] This, however, comes at the cost of ex-post consumer surplus, as indicated by a downward-sloping segment $[C, D]$. The utility's revenues are maximized for some $p^m \in (p^*, c)$, which determines the highest fixed cost for which the project remains viable (point D). While a sophisticated social planner enables a wider range of project viability (shaded-green area), there is still scope for further expansion by replacing the connection fee with a subscription plan.

A subscription plan

Instead of a connection fee F, which is paid by the present consumer, i.e., before endogenous preference formation, the social planner can use a monthly subscription fee f, which is paid *only* by the future consumer. Such a recurrent subscription fee is conceptually distinct even from our amortized representation F of the upfront connection fee; agreeing to F at time 0 commits the consumer to pay in every subsequent period irrespective of their consumption decision whereas agreeing to a subscription fee f of the same value does not entail such a commitment. Thus, a subscription plan goes beyond addressing potential credit constraints as it comes with the option of not paying the fee in the future. The social planner, aware of the time inconsistency, knows that, despite the time 0 consumer's unwillingness to pay, the time 1 consumer will be willing to pay a flow of $f = F$ beginning at time 1. Since connection fee and monthly subscription are equivalent for the utility's budget, it always (weakly) prefers to charge zero connection fee and recover its costs through the monthly subscription.[33]

[32] Revenues increase with price if the price elasticity is sufficiently low, i.e. $\alpha \leq \lambda$ (which is the case in our empirical application). Otherwise, revenues are maximized at $\{p^*, F^*\}$ and projects are not undertaken for $K > (p^* - c)e^{\lambda(\theta-\alpha p^*)} + F^*$; in terms of Figure 4, points C and D would coincide.

[33] Implementation issues with the subscription plan could arise along the transition path to the long-run steady state, which we have abstracted from in our discussion. For example, in practice,



Notes: The horizontal axis represents fixed costs per customer and the vertical axis consumer surplus. Projects on the 45-degree line generate zero net social surplus, while projects above (resp. below) generate positive (resp. negative) surplus. The vertical lines at $e^{\theta-\alpha c}$ represent the highest project fixed cost that an unsophisticated social planner can finance by setting price equal to marginal cost and A indicates the associated level of consumer surplus (blue area). Point B indicates the surplus that can be achieved by a sophisticated social planner who sets price p^* below marginal cost. This further allows financing costlier projects (green area). As the fixed cost K increases, price p^* is no longer sufficient to balance the utility's budget (point C) so that price needs to increase, with the associated decrease in consumer surplus until the project is no longer viable (point D). Efficiency is restored with a subscription plan (orange area) as all projects get financed until the fixed-cost exceeds consumer surplus (point E).

Figure 4: Two-part tariffs: connection fee vs. subscription plan

The planner's problem is given by (25), exactly as before, but is now subject to ex-post participation constraint

$$e^{\lambda(\theta-\alpha p)} \geq \alpha f. \tag{28}$$

Assuming that this constraint is non-binding, optimal markup is again $\frac{1}{\alpha}\left(\frac{1}{\lambda}-1\right)$; i.e., $p^{**}=p^{*}$, $f^{*}=F^{*}\geq K$, with equality at $\lambda=1$, and the subscription plan achieves the same consumer surplus as the connection fee *conditional* on the project being undertaken in the first place. The advantage of the subscription plan, however, is that, with it, all projects along line segment $[C,E]$ in Figure 4 will be undertaken with maximal consumer surplus (orange-shaded area).

5.2 Optimal pricing: heterogeneous preferences

With a non-degenerate preference distribution $G(.)$ over $[0,\infty)$, two-part tariffs also effectuate *between* consumer cross-subsidization (Ng and Weisser, 1974; Auerbach and Pellechio, 1978). In particular, the sophisticated planner's program, in the case of a connection fee, becomes

$$\max_{F,p} \int_{\bar{\theta}}^{\infty} e^{\lambda(\theta-\alpha p)} dG(\theta) - \alpha F[1-G(\bar{\theta})], \tag{29}$$

subject to a participation constraint that defines marginal consumer $\bar{\theta}$,

$$e^{\bar{\theta}-\alpha p} = \alpha F, \tag{30}$$

and budget constraint

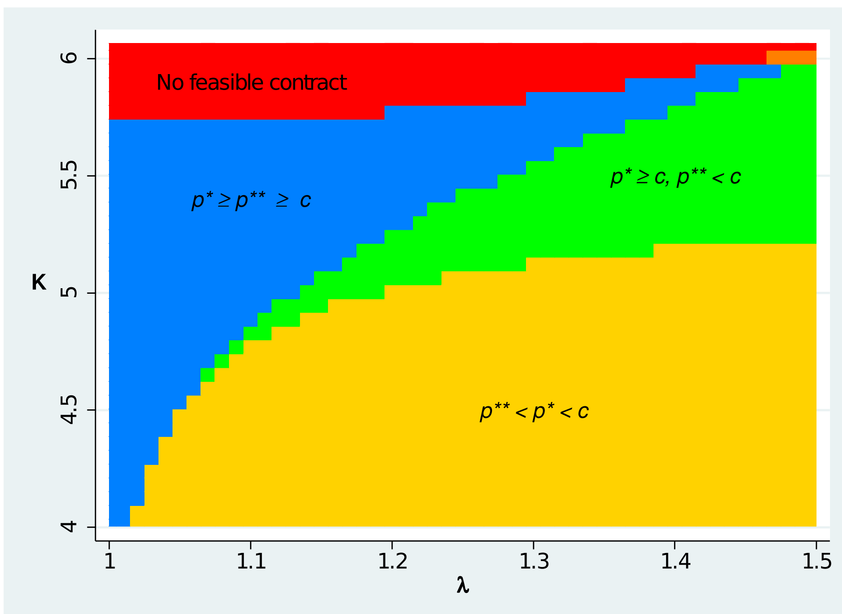
$$(p-c)\int_{\bar{\theta}}^{\infty} e^{\lambda(\theta-\alpha p)} dG(\theta) + F[1-G(\bar{\theta})] \geq K. \tag{31}$$

When $\lambda=1$, this problem reduces to Auerbach and Pellechio (1978), in which case a price $p^{*}>c$ is optimal for K sufficiently high. The markup $p^{*}-c$ acts as a tax, the greater burden of which is borne by high-θ consumers, to finance a subsidy on the

there may have to be an "introductory" period of zero or low subscription fees, which could induce households to connect that would not want to stay connected at the higher long-run subscription fee (i.e., if preferences are heterogeneous). Nevertheless, over a long enough horizon, the welfare costs of such potential "leakages" would be small.

connection fee, the main beneficiaries of which are low-θ consumers (who would not have otherwise connected). When $\lambda > 1$, however, the incentive to subsidize the price, discussed in the previous subsection, comes into play. So, preference heterogeneity and demand persistence work in opposing directions on the marginal price, with the net impact uncertain. We thus turn to numerical simulations to gain further insight.

Figure 5 illustrates optimal markup regimes for alternative combinations of per consumer fixed costs K and intertemporal dependence parameter λ.[34] When $\lambda = 1$, the optimal markups under the connection fee and subscription plans coincide and are both positive. As noted, the markup in this case serves only as a way for high-θ consumers to cross-subsidize low-θ consumers.



Notes: The graph plots solutions to the optimal connection fee with price p^* and subscription fee with price p^{**} at different fixed costs (K) and degree of intertemporal dependence (λ). In the red domain, no contract achieves cost-recovery; in the orange domain, only the subscription contract achieves cost-recovery; in the blue domain, both contracts have a positive markup; in the green domain, the connection (subscription) fee contract has a positive (negative) markup; in the gold domain, both contracts have a negative markup.

Figure 5: Markup regimes for two-part tariffs with preference heterogeneity

While a positive markup may continue to be warranted to redistribute between households (blue domain) for $\lambda > 1$, higher values of λ also imply greater deadweight loss from this taxation. Eventually, therefore, markup becomes negative, first for the

[34]We use estimates of $G(\theta)$ and α from Section 4 and fix marginal cost c at US$0.14 (see next subsection).

subscription plan (green domain) and then for the connection fee contract as well (gold domain). Markup becomes negative for smaller λ under the subscription plan because, without a connection fee, the cross-subsidization motive is attenuated. Reading Figure 5 along the vertical dimension, for a given λ, negative markups dominate (gold domain) for small K. As K increases, however, cost recovery dictates between-consumer cross-subsidization under a connection fee (green domain) and, eventually, under both a connection and subscription fee (blue domain). For high enough K, the utility, facing downward-sloping demand, cannot recover average costs (red domain).

5.3 Utility pricing scenarios

To simulate the welfare implications of alternative water price structures in our setting, we first extract the distribution of preferences for piped water in My Huong commune. Using (12), we recover $\hat{\theta}_i$ for our estimation sample. Since, as noted above, previously unconnected households are under-represented in our estimation sample, we reweight the data to ensure that the empirical distribution of $\hat{\theta}_i$ corresponds to that of the commune population.[35] Next, we parametrically estimate the distribution of preferences G using a truncated normal density fit to the empirical distribution of $\hat{\theta}_i$, where the right truncation point is the sample maximum.

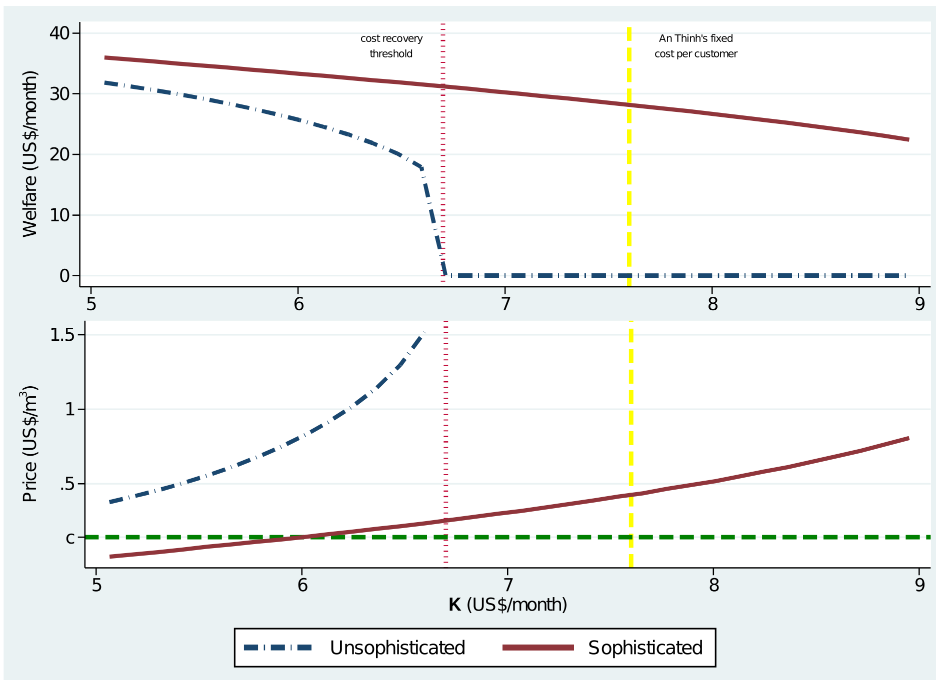
Note that $\hat{\lambda} = 2.7$ implies a substantial wedge between ex-ante and ex-post willingness-to-pay; the median value of this wedge, expressed as the ex-post to ex-ante surplus ratio, $e^{(\lambda-1)\hat{\theta}_i}$, is 3.25 with an interquartile range of 2.35-4.40. Had we instead found zero demand persistence in our pricing experiment (i.e., $\lambda = 1$), there would be no such intertemporal wedge and neither of the policies discussed next would yield any welfare gain whatsoever.

Exercise 1: Sophisticated two-part tariff

Consider the unsophisticated social planner, who errs in assuming that preferences for piped water are exogenous and unchanging ($\lambda = 1$), but who ascertains the distribution

[35] Specifically, we create an augmented representative sample by drawing 120 $\hat{\theta}_i$ (with replacement) from those of the 114 previously unconnected households included in our estimation sample; by construction, this artificial sample has 16 percent previously (as of July 2015) unconnected households as in our commune household listing.

of preferences $G(\theta)$ along with the short-run price elasticity of demand α.[36] As noted, the optimal two-part tariff (p^*, F^*) in this scenario was originally discussed by Auerbach and Pellechio (1978). Given alternative values of per customer fixed cost K, we compute ex-post social welfare at the contracts offered by the unsophisticated planner, where ex-post social welfare is the long-run ($\lambda = 2.7$) average consumer surplus net of average costs. We repeat this exercise for a sophisticated social planner, one who (correctly) assumes endogenous preferences for piped water and who knows that consumers are ex-ante unaware of having such preferences.



Notes: The top graph plots ex-post welfare (consumer surplus) under sophisticated and unsophisticated pricing against fixed cost per customer K. The bottom graph plots the corresponding marginal prices under each contract and shows the marginal cost (c) of US\$0.14/m^3. The vertical dotted pink line represents the fixed cost threshold at which the unsophisticated planner's pricing plan can no longer balance the utility's budget. Beyond this threshold, we set social welfare under the unsophisticated plan to zero. The vertical dashed yellow line represents the K reported by the An Thinh water utility.

Figure 6: Welfare with and without sophisticated pricing of new utility

Figure 6 (top panel) illustrates the benefits, in terms of US$ per month (per cus-

[36]For concreteness, suppose that this planner commissions a study in which alternative (p, F) contracts are randomly offered to potential customers and the resulting data are used to estimate take-up as a function of p, F). From the participation constraint, $e^{\theta - \alpha p} \geq F$, and a parametric assumption on G, it is easy to see that both G and α are identified. Note that this experiment is different from the one actually conducted by Lee et al. (2020) in rural Kenya inasmuch as the latter only randomizes F.

tomer), of accounting for endogenous preferences in pricing a new utility service. The welfare or consumer surplus gain from adopting a sophisticated pricing strategy, the difference between the solid and dashed curves, can be sizeable relative to average monthly household expenditures of about US$230 in our sample. The bottom panel shows that the unsophisticated planner, conditional on operating the utility at all, charges a much higher price for water than the sophisticated planner, and one well in excess of marginal cost; indeed, for fixed cost per consumer below US$6, the sophisticated planner would set a sub-marginal cost price.[37]

To pin down the relevant scenario, we obtained actual fixed cost figures from An Thinh water utility, consisting of their initial investment in plant and equipment (including the underground pipe network for the three villages of the commune), annual operating expenses, as well as the cost of a home connection and water meter, all together amounting to US$7.6 per month per customer.[38] At $K = 7.6$ in Figure 6, we see that the social value of the utility when decisions are taken by an unsophisticated social planner is zero. From this erroneous perspective, there is no combination of markup and connection fee that would allow the utility to cover its costs and, therefore, the project would not be undertaken in the first place. By contrast, a sophisticated social planner, who is attentive to endogenous preferences, would undertake the project, break even, and produce a positive social surplus of around US$28 per customer per month. This would be achieved by charging a price of US$0.42 per m^3 and a connection fee equivalent to a monthly flow of US$6.3; in all, the average consumer would spend roughly 4 percent of their income on piped water (with none of their tax revenue diverted to subsidize the utility). While the welfare gain of US$28 per month is relatively large, note that it reflects a change in consumption along the extensive margin, since zero piped water is consumed under unsophisticated pricing.[39] At any rate, this finding

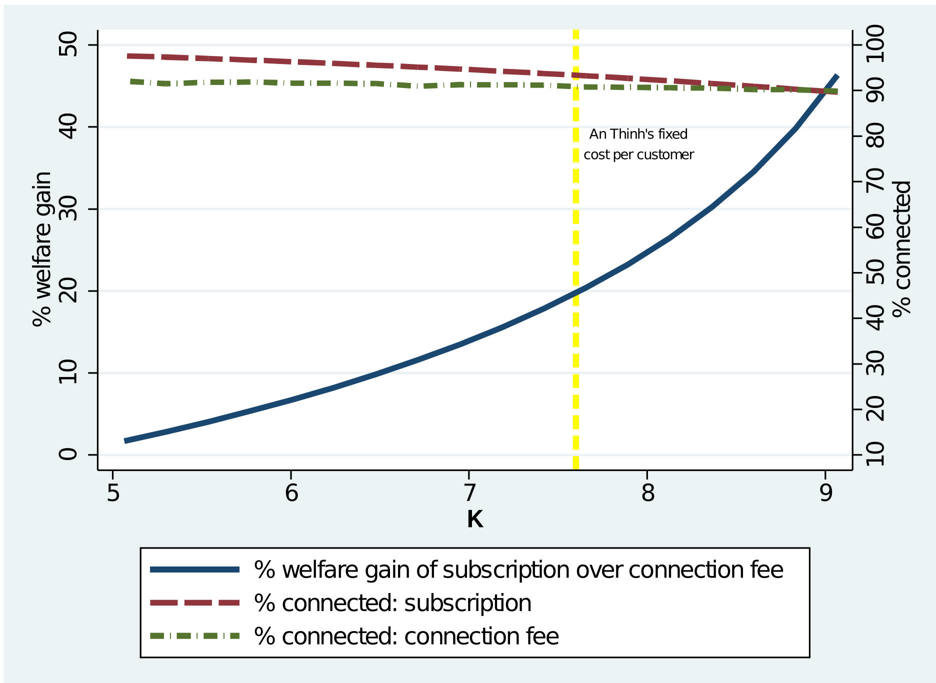
[37]We assume that $c =$US$0.14 per m^3, which is the marginal cost reported to us by An Thinh water utility (primarily fuel costs for pumping water from the river and to its customers).

[38] Specifically, initial investment was reported to us as US$9.7 million in current dollars, annual operating costs at US$30,000, and household connection cost at US$84 per customer. We convert all of these figures to monthly flows per customer at the plant's reported capacity of 10,000 customers. In the case of plant and equipment, we do the stock to flow conversion using a 5 percent real interest rate, equal to the lending rate minus the rate of inflation (both for 2019) as reported for Vietnam in World Bank statistics. Since the utility is privately owned and run as a business (albeit subject to government rate regulation), the US$7.6 figure probably reflects the social opportunity cost of piped water provision in rural Vietnam quite accurately.

[39]At $K = 6$, for instance, the implied change in piped water usage would be 'marginal' and, hence,

suggests caution in relying *solely* on assessments of consumer demand or willingness-to-pay made before consumers have ever connected to judge the economic viability of a utility service.

Exercise 2: Subscription plan



Notes: The graph plots the percent welfare gain from switching from a connection fee to a subscription plan (solid blue) and the connection rates under each. The vertical dashed yellow line represents the K reported by the An Thinh water utility.

Figure 7: Subscription fee versus connection fee contract

We next turn to the best way to price the utility service for a sophisticated planner. Figure 7 (solid blue line) shows the percent welfare gain, as a function of per customer fixed cost K, from switching away from the optimal two-part tariff with a connection fee to a deferred subscription plan. Depending on the level of costs that the utility needs to cover, the benefits from moving to a subscription plan can also be substantial. As the long-dashed green and dash-dotted orange lines indicate, while a subscription plan allows higher participation, the difference in connection rates is, at most, around 5 percentage points. The bulk of the welfare gains thus come at the intensive margin

the welfare gain in Figure 6 would only be about US$7 per month.

rather than at the connection margin, which is to say as a result of the lower markup that a subscription plan entails.[40]

In the case of An Thinh water utility, reading off the dashed vertical (yellow) line in Figure 7, there would be a 20 percent welfare gain in moving from a (prorated) connection fee to the subscription plan, equivalent to US$5.6 worth of consumer surplus per month on average.

6 Conclusion

New residential utilities are one of the mileposts of economic development, and yet their pricing and economics of provision have not heretofore been studied in the (plausible) case where preferences for the service evolve endogenously over time. To do so, we generated experimental variation in the price of residential piped water in a setting where households were still transitioning away from a traditional mode of water delivery. Consistent with the new piped water service being an experience good, being habit-forming, or being some combination of the two, we find short-term persistence in demand: high consumption today increases consumption in subsequent months.

Insofar as it is not internalized by consumers, such positive intertemporal dependence alters the calculus of two-part tariffs in two ways, first by providing an incentive to subsidize consumption with a lower marginal price, possibly one even lower than marginal cost, and second by providing an incentive to defer lump-sum payments until long-run preferences are formed. We have shown in our particular setting that, whereas long-run demand for the new utility is such that it can fully cover its costs through a connection fee and markup, it cannot do so if it prices its service without attention to endogenous preferences. We further find that irrespective of credit constraints, a subscription approach to deferring payment would achieve a 20 percent welfare gain over the corresponding optimal contract with an upfront connection fee.

Neither optimal pricing plan that we consider in this paper is without risk to the utility, given imperfect knowledge of consumer preferences. On the one hand, if the

[40] In all the simulations, we impose $p \geq 0$, given the potential practical issues involved in charging a negative water price. Recall from the lower panel of Figure 6 that the connection fee contract has $0 < p < c$ for $K < 6$ and $p \geq c$ thereafter; thus $p \geq 0$ never binds. For the subscription contract, however, this constraint binds for all values of K shown.

utility overestimates consumption persistence and ultimately charges consumers too high a subscription fee, say, then it may not be able to cover its costs ex-post, resulting in socially wasteful investment. On the other hand, if the utility underestimates persistence and, as a result, finds that it could never recoup its costs through a subscription plan, then it may choose not to operate at all, forgoing a socially desirable project. Understanding the tradeoff between these two types of welfare losses is left as a topic for future research.[41]

Appendix material

Appendix material for this paper is available at *Review of Economic Studies* online

Data Availability Statement

The data and code underlying this research are available on Zenodo at `https://doi.org/10.5281/zenodo.10965745`

References

Acland, Dan and Matthew R Levy, "Naiveté, projection bias, and habit formation in gym attendance," *Management Science*, 2015, *61* (1), 146–160.

Ahrens, Steffen, Inske Pirschel, and Dennis J Snower, "A theory of price adjustment under loss aversion," *Journal of Economic Behavior & Organization*, 2017, *134*, 78–95.

Allcott, Hunt and Todd Rogers, "The Short-Run and Long-Run Effects of Behavioral Interventions: Experimental Evidence from Energy Conservation," *American Economic Review*, October 2014, *104* (10), 3003–37.

Alvarez, Fernando E and David O Argente, "Consumer surplus of alternative payment methods: Paying uber with cash," Technical Report, National Bureau of Economic Research 2020.

[41]One can speculate that the Government of Vietnam's piped water provision policy, which involves capital subsidies to providers coupled with provincial level price setting, is one mechanism to pool precisely this type of risk.

Andrews, Donald WK and Xu Cheng, "Estimation and inference with weak, semi-strong, and strong identification," *Econometrica*, 2012, *80* (5), 2153–2211.

Auerbach, Alan J and Anthony J Pellechio, "The two-part tariff and voluntary market participation," *The Quarterly Journal of Economics*, November 1978, *92* (4), 571–587.

Augenblick, Ned and Matthew Rabin, "An experiment on time preference and misprediction in unpleasant tasks," *Review of Economic Studies*, 2019, *86* (3), 941–975.

Becker, Gary S. and Kevin M. Murphy, "A Theory of Rational Addiction," *Journal of Political Economy*, 1988, *96* (4), 675–700.

Belloni, A., D. Chen, V. Chernozhukov, and C. Hansen, "Sparse Models and Methods for Optimal Instruments With an Application to Eminent Domain," *Econometrica*, 2012, *80* (6), 2369–2429.

Bergemann, Dirk and Juuso Välimäki, "Dynamic pricing of new experience goods," *Journal of Political Economy*, 2006, *114* (4), 713–743.

Berkouwer, Susanna and Joshua T. Dean, "Credit, Attention, and Externalities in the adoption of Energy Efficient Technologies by Low-Income Households," Manuscript, UC Berkeley 2021.

Costa, Francisco and Francois Gerard, "Hysteresis and the Welfare Effect of Corrective Policies: Theory and Evidence from an Energy-Saving Program," Working Paper 24608, National Bureau of Economic Research May 2018.

Deryugina, Tatyana, Alexander MacKay, and Julian Reif, "The long-run dynamics of electricity demand: Evidence from municipal aggregation," *American Economic Journal: Applied Economics*, 2020, *12* (1), 86–114.

Devoto, Florencia, Esther Duflo, Pascaline Dupas, William Parienté, and Vincent Pons, "Happiness on tap: Piped water adoption in urban Morocco," *American Economic Journal: Economic Policy*, 2012, *4* (4), 68–99.

_ , _ , _ , **William Parient, and Vincent Pons**, "Happiness on Tap: Piped Water Adoption in Urban Morocco," *American Economic Journal: Economic Policy*, September 2012, *4* (4), 68–99.

Diakité, Daouda, Aggey Semenov, and Alban Thomas, "A proposal for social pricing of water supply in Côte d'Ivoire," *Journal of Development Economics*, 2009, *88* (2), 258–268.

Dupas, Pascaline, "Short-run subsidies and long-run adoption of new health products: Evidence from a field experiment," *Econometrica*, 2014, *82* (1), 197–228.

Fobi, Simone, Varun Deshpande, Samson Ondiek, Vijay Modi, and Jay Taneja, "A longitudinal study of electricity consumption growth in Kenya," *Energy Policy*, 2018, *123*, 569–578.

Hansen, Bruce, *Econometrics*, Princeton University Press, 2022.

Hansen, Bruce E, "Inference when a nuisance parameter is not identified under the null hypothesis," *Econometrica: Journal of the econometric society*, 1996, pp. 413–430.

Horowitz, Joel L, "Bootstrap methods in econometrics: theory and numerical performance," *Econometric Society Monographs*, 1997, *28*, 188–222.

Hussam, Reshmaan, Atonu Rabbani, Giovanni Reggiani, and Natalia Rigol, "Habit Formation and Rational Addiction: A Field Experiment in Handwashing," Harvard Business School Working Papers 18-030, Harvard Business School September 2017.

Ito, Koichiro, "Do consumers respond to marginal or average price? Evidence from nonlinear electricity pricing," *The American Economic Review*, 2014, *104* (2), 537–563.

_ **and Shuang Zhang**, "Reforming Inefficient Energy Pricing: Evidence from China," Working Paper 26853, National Bureau of Economic Research March 2020.

— , **Takanori Ida, and Makoto Tanaka**, "Moral suasion and economic incentives: Field experimental evidence from energy demand," *American Economic Journal: Economic Policy*, 2018, *10* (1), 240–67.

Jessoe, Katrina and David Rapson, "Knowledge is (less) power: Experimental evidence from residential energy use," *American Economic Review*, 2014, *104* (4), 1417–38.

Joskow, Paul L, "Regulation of natural monopoly," *Handbook of law and economics*, 2007, *2*, 1227–1348.

Kilian, Lutz, "Accounting for lag order uncertainty in autoregressions: the endogenous lag order bootstrap algorithm," *Journal of Time Series Analysis*, 1998, *19* (5), 531–548.

Kőszegi, Botond and Matthew Rabin, "A model of reference-dependent preferences," *The Quarterly Journal of Economics*, 2006, *121* (4), 1133–1165.

Laffont, Jean-Jacques and Jean Tirole, *A theory of incentives in procurement and regulation*, MIT press, 1993.

Lee, Kenneth, Edward Miguel, and Catherine Wolfram, "Experimental evidence on the economics of rural electrification," *Journal of Political Economy*, 2020, *128* (4), 1523–1565.

Loewenstein, George, Ted O'Donoghue, and Matthew Rabin, "Projection bias in predicting future utility," *The Quarterly Journal of economics*, 2003, *118* (4), 1209–1248.

McRae, Shaun, "Infrastructure quality and the subsidy trap," *The American Economic Review*, 2015, *105* (1), 35–66.

Nelson, Phillip, "Information and consumer behavior," *Journal of political economy*, 1970, *78* (2), 311–329.

Ng, Yew-Kwang and Mendel Weisser, "Optimal pricing with a budget constraint–The case of the two-part tariff," *The Review of Economic Studies*, 1974, *41* (3), 337–345.

Pollak, Robert A., "Habit Formation and Dynamic Demand Functions," *Journal of Political Economy*, 1970, *78* (4), 745–763.

Reiss, Peter C and Matthew W White, "Household electricity demand, revisited," *The Review of Economic Studies*, July 2005, *72* (3), 853–883.

Shapiro, Carl, "Optimal pricing of experience goods," *The Bell Journal of Economics*, 1983, pp. 497–507.

Szabó, Andrea, "The Value of Free Water: Analyzing South Africa's Free Basic Water Policy," *Econometrica*, 2015, *83* (5), 1913–1961.

Taylor, Vanessa and Frank Trentmann, "Liquid politics: water and the politics of everyday life in the modern city," *Past and Present*, 2011, *211* (1), 199–241.

Thaler, Richard, "Mental accounting and consumer choice," *Marketing science*, 1985, *4* (3), 199–214.

Wilson, Robert, *Nonlinear pricing*, Oxford University Press on Demand, 1993.